2022 ACTA UNIVERSITATIS CAROLINAE PAG. 35–50
PHILOLOGICA 1 / PHONETICA PRAGENSIA

WORD-BOUNDARY GLOTTALIZATION AND ITS COGNITIVE ASPECTS: A REACTION-TIME STUDY

MICHAELA SVOBODA, PAVEL ŠTURM

ABSTRACT

This paper deals with the perceptual impact of using word-initial glottalization as a boundary signal. Research into glottalization is discussed before presenting a reaction time (RT) experiment based on a word-monitoring paradigm that aims to investigate the cognitive effect of the presence of glottalization in native Czech listeners' perception. Natural and acoustically manipulated sentences selected from a spontaneous political debate were used as stimuli. The low level of control of the material makes the design relatively innovative in comparison with similar previous studies. Fifty listeners were instructed to press a button as soon as they heard a specified target word embedded in a sentence, where a preceding carrier word included glottalization or linking. The results support the hypothesis that the presence of glottalization facilitates the processing of subsequent words, but the temporal scope of the effect varied. The experiment also raises new methodological issues and offers suggestions for further RT research.

Key words: cognition, Czech, glottalization, perception, reaction time

1. Form and function of glottalization

This paper examines 'glottalization', a specific phonetic event originating in the larynx that has been given a variety of labels. Since terminological inconsistencies and idiosyncratic interpretations can generate only confusion in the field, we shall start with defining and clarifying the relevant terms, for which differentiating between form and function will be a key prerequisite.

In the narrow, articulatory sense, the GLOTTAL/LARYNGEAL STOP (GLOTTAL/LARYNGEAL PLOSIVE) is a non-continuant obstruent articulation at the glottis, i.e., a complete closure at the adducted vocal folds. The air is compressed in front of the obstruction until the closure is released, resulting in a short burst or 'plosion'. In this perspective, the glottal stop is just one of the items in the IPA row of plosives, transcribed as [ʔ]. Acoustically, glottal plosives are not all identical, as they differ in the shape of the waveform (see Skarnitzl, 2004 for details and images). What is common to them is a percept of a sudden change – cessation to silence or rise from silence (Cruttenden, 2014: 182).

https://doi.org/10.14712/24646830.2022.27

© 2022 The Author. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0).

However, in identical positions, in free variation with the glottal plosives, other speech sounds may occur. Although they are quite different from plosives acoustically, they are quite similar perceptually (namely, a break between the two neighbouring sounds is perceived). The glottal plosive may be seen as an extreme position on the phonation continuum (Gordon & Ladefoged, 2001: 384), in which the non-modal glottal phenomena between modal phonation and glottal closure are perceptually equivalent. In fact, in genuine spoken interactions, complete glottal closures are less common than various lenited variants, including creaky voice or laryngealization characterized chiefly by irregularity in vibration and its lower frequency (Ashby & Przedlacka, 2011, 2014, for English). Several publications deal with the acoustics of these glottal events (Docherty & Foulkes, 1999; Redi & Shattuck-Hufnagel, 2001; Skarnitzl 2004; Keating, Garellek & Kreiman, 2015).

All the manifestations mentioned above can be subsumed under **one common term**, namely glottalization. Any such glottal activity – glottal plosive, creaky voice – would be perceived as a single phenomenon – glottalization. Sometimes, the term glottal stop is used in this broader sense as well. We will be referring to 'glottalization' in the rest of the paper. However, not every occurrence of creaky phonation (laryngealization) is an instance of glottalization. For example, creaky phonation may be associated with ends of prosodic phrases (Local, Wells & Sebba, 1985; Ogden, 2001), be a habitual (sociolinguistic) feature of one's voice (Yuasa, 2010; Wolk, Abdelli-Beruh & Slavin, 2012) or can signal a pragmatic or affective meaning (Gobl & Ní Chasaide, 2003). This type of creak seems to be prosodically grounded compared to the segmentally defined creak in glottalization.

Finally, it is important to establish the **function of glottalization** in the given language. As a functionally defined unit, glottalization can be employed in two main ways. It can be used (1) contrastively, where [ʔ] is in opposition to realizations of other consonant phonemes. In languages such as Hawaiian or Arabic, it is the primary allophone of the phoneme (Ladefoged & Maddieson, 1996; Maddieson, 1984), while in languages such as English or the Philippine language Ilokano, it is a regional or stylistic variant (Cruttenden, 2014; Olaya, 1967). Specifically, the oral plosive [t] (*cat, later*) in one accent of English may correspond to [ʔ] (glottalization) in other accents, or it may be conditioned by the environment or speaking style within one accent (see e.g., Fabricius, 2002; Gavaldà, 2016). Moreover, glottalization is often used as (2) a boundary signal of a lexical/grammatical unit, i.e., it has a demarcative function. Czech or Polish are typical examples of languages where the presence of glottalization cues morphosyntactic word boundaries when the word starts with a vowel ([ʔ]*akustika*, 'acoustics'), and some internal morpheme boundaries (*na*[ʔ]*opak*, 'on the contrary'). The two functions are not incompatible, as some languages (e.g., English) have both contrastive and demarcative glottalization.

An anonymous reviewer pointed out another attested function, (3) hiatus breaking, in which sequences of two vowels are split by glottalization as a sort of syllable reorganization or onset formation (for instance in German *Theater* [teˈʔaːtɐ], 'theatre' or Czech *teoreticky* [ˈtɛʔorɛtɪt͡skɪ] 'theoretically'). However, the use of such forms needs to be examined empirically.

2. Use of boundary glottalization in languages

This paper deals with glottalization in its demarcative function, i.e., as a boundary cue. Available sources suggest that its fulfilment is not a binary matter; rather, there is a continuum between the tendency to glottalize initial vowels and the opposite tendency to link initial vowels to the previous words. The studies presented below offer evidence of substantial difficulty in trying to establish an 'inherent' glottalization rate for each language, as it seems that **context (speaking situation)** is a key factor. Therefore, it is imperative to consider the type of the examined material, especially its position on the scale of spontaneity or level of control (Wagner, Trouvain & Zimmerer, 2015). In cross-linguistic comparison, one cannot confront for instance the material of read speech with a spontaneous dialogue.

Nevertheless, owing to considerable research interest in the topic, a certain picture on the use of glottalization emerges at least for some languages. For instance, German is often characterized as a language that employs glottalization quite extensively. Kohler (1994) measured the rate of glottalization in read speech from the Phonetic Data Bank of German, yielding 79% of words with initial vowels being glottalized. This study also suggested that a higher glottalization rate can be expected in stressed syllables, which was confirmed in subsequent research (Pompino-Marschall & Żygis, 2010; Malisz, Żygis & Pompino-Marschall, 2013). Moreover, these later studies were of somewhat lower level of control, represented by prepared speech, and established, respectively, 50–70% and 63% of glottalized initial vowels. Other contributing factors were increased speech tempo, leading to higher glottalization rates, or the semantic status of the word, with lexical words being glottalized more often than grammatical words.

Moving to the opposite tendency, English is a language that uses linking rather than boundary glottalization (Cruttenden, 2014). However, this does not mean that we should expect no glottalization at word edges at all (see previous section for contrastive glottalization). For instance, Dilley, Shattuck-Hufnagel and Ostendorf (1996) showed that news-reading was associated with substantial variability across speakers, ranging from 13% to 44%. Another finding was the contribution of prosodic factors, including lexical stress (as in German) or the strength of the prosodic boundary (stronger boundaries tended to involve more glottalization). Many Romance languages – French, Spanish, Italian, Portuguese – behave similarly to English in this respect (Skákal, 2013; Di Napoli, 2015; Skarnitzl, Čermák, Šturm, Obstová & Hricsina, 2021). The main strategy in these languages is to link words smoothly without glottalization (thus *por‿otro, que‿ahora* and not *por*[ʔ]*otro* etc.).

The research on Czech examined several factors and used material with various levels of control. Veroňková (2016) confirmed that the requirement to use glottalization after non-syllabic prepositions (Hála, 1967) was to a large extent obeyed in read speech. Volín (2012) brought a comparison of professional news-reading with semi-spontaneous dialogues. Glottalization rate was high in the former speaking style (97% by female newsreaders, 88% by their male colleagues) and much lower in the latter situation (men in spontaneous dialogues glottalized only 41% of the potential cases, women 65%). The results thus suggest an effect of material/context on the one hand, and gender on the other (compare similar effects of gender in Dilley et al., 1996;

Skákal, 2015; Gráczi & Markó, 2018; Kopečková, 2020). Further factors include the segmental and prosodic environment. Palková (2016), examining the read speech of students, compared V#V and C#V sequences in various positions within a prosodic phrase. Although the segmental environment did not play a significant role (similarly to Veroňková, 2016), strength of the prosodic boundary was a crucial factor (see also Dilley et al., 1996).

3. Perception of boundary glottalization

To fully understand a sound pattern such as glottalization, its perceptual impact must be thoroughly examined. The use of glottalization can be experimentally treated from several perspectives but for reasons of scope we will limit ourselves to discussing the effect of glottalization on the online processing of speech, namely on how words are recognized in an utterance.

A comparison of a glottalized sequence with a non-glottalized (linked) sequence is necessary. If glottalization plays an important role in processing the stream of speech, the former should have a processing advantage over the latter. Two situations should be distinguished. On the one hand, a vowel-initial word is cued with glottalization, thus comparing for instance *stál_a čekal* with *stál* [ʔ]*a čekal* ('He stood there and waited.'). On the other hand, two actual words differ solely in the presence of glottalization, such as *kočku* ('cat') × *k*[ʔ]*očku* ('to a loop'), *sokem* ('enemy') × *s*[ʔ]*okem* ('with an eye'), *kůlu* ('pole') × *k*[ʔ]*úlu* ('to a hive'), so that the use of glottalization would be crucial to understanding. However, it is remarkably difficult to find such oppositions. Moreover, the grammatical and pragmatic environment would make it even more difficult to provide a source of confusion as to which member of the pair is present. The context is sufficient to disambiguate the variants in a sentence.

However, it would be wrong to conclude that using or not using glottalization is entirely irrelevant to the online processing of speech. The brain operates as a predictive device that constantly lowers the amount of uncertainty about the upcoming events (Grossberg, 2003). As the signal unfolds, the brain matches the auditory encoding of the current acoustic signal to a set of stored representations. Lexical access activates a number of representations ('words') based on past experience (e.g., context, topic, a closed × open set of answers to a question) and on the auditory signal. For instance, hearing [kot] might activate the words *kotel*, *kotva*, *kotouč*, but also *k odkazu* [kotkazu], *k odpovědi* [kotpovjɛɟɪ] and so on. In contrast, hearing [kʔot] might skip the irrelevant words, as *kotel*, *kotva*, *kotouč* do not include the glottal stop in their representation (the speaker has never heard them with glottalization, so has not stored it as such), and only the words starting with /ot/ after the preposition *k* will be the basis of assembling the candidate list. Although the brain will eventually solve the segmentation problem in both cases, [kot] and [kʔot], not least because of the contextual disambiguation, it is safe to assume that **the presence of glottalization provides a processing benefit** and is cognitively less demanding than the non-glottalized variants.

To our knowledge, only three experiments have been designed to test this prediction explicitly (Bissiri, Lecumberri, Cooke & Volín, 2011; Volín, Uhrinová & Skarnitzl,

2012; Schwartz, Rojczyk & Balas, 2015). They all employ a word-monitoring paradigm, in which a subject's response to a specified target word is collected, measuring the reaction time (RT) of the response. The material was read English, but the listeners differed in their native languages and their typical glottalization rates. Bissiri et al. (2011) compared Czech, Spanish and native English listeners, Volín et al. (2012) Czech, Slovak and native English listeners, and Schwartz et al. (2015) used Polish listeners of varying English experience. The hypothesis was that the presence of glottalization would decrease the RT to that word compared to the linking condition. The two early studies in addition supposed that Czech listeners would benefit more from the presence of glottalization than English, Spanish or Slovak listeners, who use glottalization less extensively. Surprisingly, all groups were associated with decreased RTs in response to glottalization (despite some differences). The Polish experiments found an effect in the predicted direction, but it was small in size and also not much reliable.

However, the Czech studies differed from the Polish study in several important respects. Schwartz et al. (2015) included mostly predictable short sentences (*Bob ate the whole chicken*.), and the occurrence of glottalization or linking was always natural (produced by a real speaker in that context). In contrast, the material for Bissiri et al. (2011) and Volín et al. (2012) was news-reading, and the sentences started in medias res, with the beginning cut off so that the sentence would be unpredictable (*with ten men after the striker Thierry Henry*; the word *after* served as a target). Moreover, each sentence was presented in two conditions, with or without glottalization, which was done by means of manipulation of the speech signal (adding or removing glottalization), resulting in pairs of stimuli differing solely in that respect.

4. Experiment with RT measurement

The studies mentioned above examined the effect of glottalization in L2 English on non-native listeners. Whereas Czech listeners seem to benefit from the presence of initial glottalization in English, it is also necessary to determine the perceptual impact of boundary glottalization using native Czech material. The current experiment aims to fill that void. Our second goal is to move beyond scripted speech and use material with a lower level of control (Wagner et al., 2015), namely a political TV debate.

4.1 Method

4.1.1 Material

The recordings come from the political discussion TV programme *Nedělní partie*, in which a moderator (male, 31 years) hosts two guests (in our case, both male, 39 and 55 years) who present their views and argue with the moderator or with each other. The style is therefore semi-formal, the type of material can be labelled as non-scripted speech (Wagner et al., 2015) rather than truly spontaneous speech. The particular 50-minute episode was aired on TV Prima in the Czech Republic on 25/4/2010. The debate reflects some of the important political topics of the time.

The debate comprises 5841 orthographic words. Of these, 650 words (11%) begin with an initial vowel, which is typical for Czech initial syllables (12.2% according to Šturm & Bičan, 2021: 369). We conducted a simple acoustic analysis of the debate to determine the rate of boundary glottalization of the three speakers (word-internal glottalization was not examined). Tokens that were preceded by a pause or where several speakers spoke simultaneously were ignored. The host produced glottalization in 38%, and the two guests in 44% and 54% of the possible contexts. Such rates seem to be quite low, but it needs to be stressed that previous reports (e.g., Palková, 2016; Veroňková, 2016; partly Volín, 2012) were based on read material.

4.1.2 Stimuli for the perception test

Due to the low level of control of the material, we had no means of determining or shaping the sentences; it was only possible to select suitable items from the corpus. Such recordings (carrier + prompt embedded in sentences or their parts) had to be 'clean' (i.e., with no hesitations, interruptions, overlaps, or external noise), the site of glottalization should be suitable for unnoticed manipulation (see below), and the prompt should not have strong collocations with the preceding words, as this could weaken any perceptual effect.

The resulting set of 40 items (see the Appendix) is necessarily – but also intentionally – heterogenous. Twenty items occurred with glottalization, twenty without glottalization. Frequencies of the words, semantics, stress position, tempo or other factors were not controlled. This should not pose a problem, as we used planned paired comparisons between identical sentences where one version includes glottalization, while the other does not. Therefore, any such effect should be constant within each text. Importantly, the position of the prompt varied, and the beginning of the items does not correspond to sentence beginnings.

The manipulation of the carrier word was performed in Praat (Boersma & Weenink, 2021). The process differed depending on the sequence (C#V, V#V, types of V) and the type of manipulation (adding/removing glottalization). Adding glottalization was relatively easy: a glottal stop or its equivalent was copied from a suitable context from a different part of the debate, replacing the original linked context. Removing glottalization was more difficult. Again, a similar sequence was found in the debate – one without glottalization – which replaced the original glottalized token, or the glottalization interval was deleted. The neighbouring vowels usually had to be compensated in duration and adjusted (shortened or lengthened) manually or by PSOLA manipulations.

One of the key aspects that differentiates our experiment from those mentioned in Section 3 is the structure of the stimuli. In the previous setups, the target word or (i.e., the prompt, which is monitored for by the listeners) was at the same time a carrier word (i.e., the site of glottalization). The word *after* – with or without glottalization – was therefore printed on screen and reacted to in Volín et al.'s (2012) experiment. The current experiment involves a sequence of the manipulated carrier word followed by the target word (prompt), with varying distance of 1 to 7 syllables between the initial syllables. Therefore, the carrier is vowel-initial, but the target word does not start with a vowel (see Figure 1).

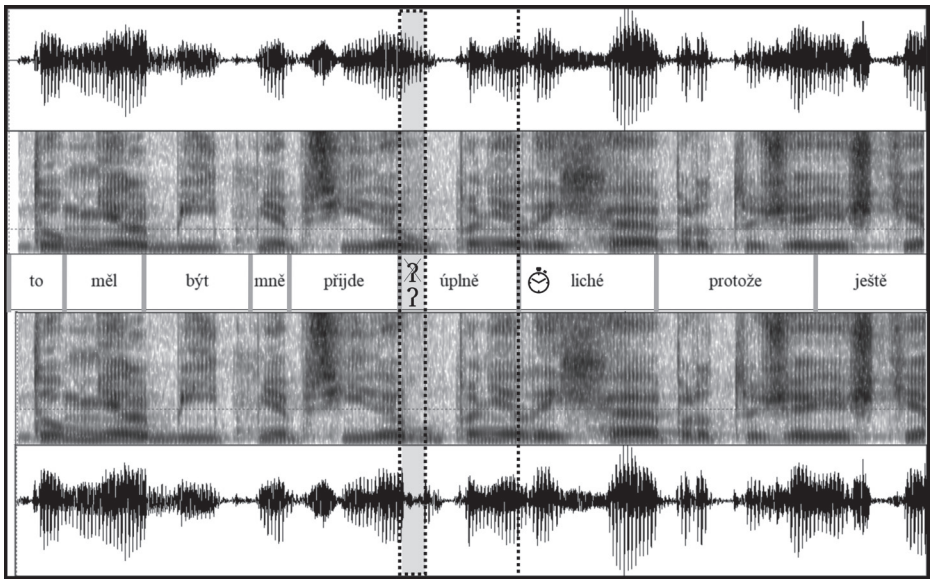

Figure 1. Example of a stimulus pair. The word *liché* is the prompt (target) displayed on the screen. The word *úplně* is the carrier – at its beginning, there is either linking (top) or glottalization (bottom). RT is measured from the start of the prompt (dotted line with a clock symbol). The prompt here is three syllables away from the glottalization site.

Such a change in the design of the experiment was done for the following reasons. First, it is not clear from which point RT should be measured in a vowel-initial word, as glottalization interferes with the temporal structure of the vowel. Having the onset of measurement on the following word, which is identical in both cases, solves this dilemma. Second, it is reasonable to assume that the effect of glottalization appears with a delay, i.e., the processing of the following word(s) would be affected. Finally, the modification also reflects our interest in knowing whether the distance between the carrier and prompt plays an important role. In some items, the glottalized syllable immediately precedes the prompt, in others it precedes it at a distance.

Two sets of 20 stimuli were originally produced with or without glottalization (Appendix). Two sets of 20 new stimuli were created by manipulation (adding or removing glottalization from the previous two sets). Therefore, 2 × 40 target items were used. In addition, 32 other items (almost 30% of the total number) served as fillers and distractors, in which the prompt occurred very early or late in the sentence, or it did not occur at all. Inclusion of such trials makes the task of recognizing a word more interesting, less monotonous and, above all, unpredictable to the listener. Finally, six other items were selected for a training session; their structure covered all the types used (target, empty, early, late). In sum, 118 stimuli were presented to the participants, who needed from 15 to 19 minutes to complete the session.

4.1.3 Listeners and procedure

The experiment was administered to 50 native speakers of Czech (18–53 years, mean = 24 years). There were 30 women and 20 men, and all but three came from Prague or Central Bohemia, most of them students of various universities.

The sessions took place in 2021 at the Institute of Phonetics, Charles University, Prague, in a sound-treated room without any disturbances or interruptions. The experiment was run in *Dmdx* (Forster & Forster, 2003; see Šturm & Volín, 2012). The sound level was first adjusted according to the participant's needs. The experiment started with written instructions specifying the task and procedure. Each item consisted of the following: (1) the prompt, displayed in the centre of the screen in uppercase letters; (2) a short audio signal (beep) followed by silence, indicating the start of the stimulus; (3) the stimulus itself. The participants were instructed to press a large button in front of them as soon as they heard and recognized the word presented orthographically (see Kilborn & Moss, 1996, for a methodological discussion of the task). The button (Black Box Toolkit) was designed specifically for RT measurements with millisecond accuracy. When the sentence did not include the target word, the participant just waited. No other action was needed, and the next item appeared automatically.

After a training session, the 112 remaining stimuli were presented in four blocks (with a short, one-minute break between them). Each block contained 20 target items and 8 fillers. The order of items within a block was randomized for each participant. However, the order of blocks was semi-random due to the fact that pairs of stimuli (glottalization vs. linking) were used. Therefore, a glottalized version of a carrier could not appear in the same half as the counterpart version (in other words, the same text was restricted to different halves). To prevent any order effects, two versions of the experiment were created, balanced across participants. In one, blocks A and B (in random order) were followed by blocks C and D (in random order). In the other version, blocks C and D preceded blocks A and B.

4.1.4 Data processing and analysis

There were 4000 target observations (without fillers). As is common practice in RT research, two types of reactions were dismissed: anticipations (button is pressed sooner than the brain's cognition allows for a proper response) and misses/delays (button is pressed later than what is regarded as immediate reaction). It is clear that defining a valid response is not a simple task, and the decision differs across studies (see Jiang, 2012). In line with Bissiri et al. (2011), we defined a 'hit' as a response between 150 ms and 1000 ms. As a result, 3617 responses remained in the dataset. Subsequently, unreliable listeners and items were discarded (for simple tasks, Jiang, 2012 recommends dismissing listeners with an error rate over 20%). Two listeners were discarded (error rate 23%), and three pairs of items (error rate 80–90%; the remaining items had error rates below 13%). The final dataset comprised 3466 observations.

As is typical in RT research, histograms revealed a highly positive skew in the RT distribution. Therefore, the RT values in milliseconds were logarithmically transformed, which resulted in a normal distribution.

Statistics and data visualization was performed in *R* (R Core Team, 2020), using the libraries *lme4* (Bates, Mächler, Bolker & Walker, 2015), *emmeans* (Lenth, 2022) and *ggplot2* (Wickham, 2009). LME models were created separately for the two types of manipulation, with glottalization (linked × glottalized) and distance (1, 2, 3, 4, 5-7 syllables) as fixed effects and listener and prompt as random effects. The distance was a factor variable, not numeric.

4.2 Results

A simple comparison of median values for glottalized and non-glottalized conditions (Fig. 2) revealed a negligible difference of 0.008 on the logarithmic scale (translating to approx. 3 ms after back transformation, or less than 1%). Since the two types of manipulation could yield different results, Figure 3 splits the dataset according to whether glottalization was added or eliminated. However, very similar values appeared in each condition (a change of 3 ms, or 1%, for addition, a change of 10 ms, or 3%, for elimination). The violin plots also confirm the normal distribution of log-transformed values.

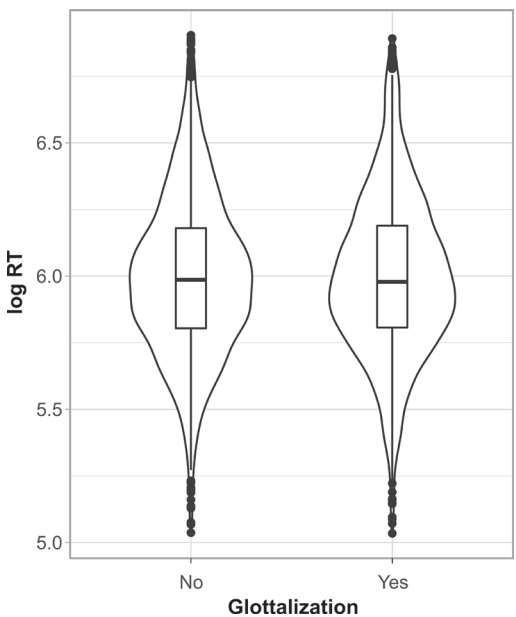

Figure 2. Logarithmic RTs of glottalized and linked stimuli in the full dataset.

Although glottalization does not seem to elicit any substantial RT difference in the participants' behaviour, we must still consider the effect of the distance between the prompt and the carrier. Moreover, the graphs so far described independent data. Figure 4 now displays paired differences between the two versions of each stimulus, where positive values mean shorter RTs in the glottalized version. Only boxplots are shown for better

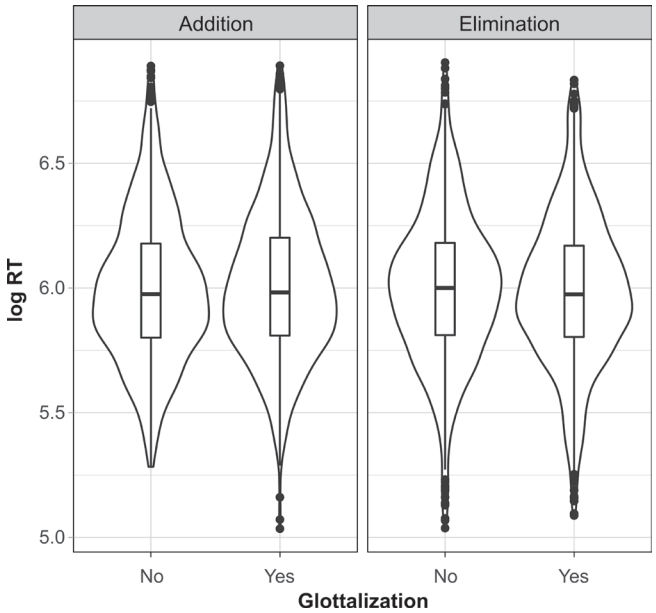

Figure 3. Logarithmic RTs of glottalized and linked stimuli in two types of manipulation. Note that "No" under Addition and "Yes" under Elimination mark the original as compared to the manipulated stimulus.

clarity. The graph on the left suggests that addition had little impact on the RTs, and with distances of 3 and 4 syllables even seemed to produce the opposite effect (higher RTs of glottalized stimuli). In contrast, elimination yielded either small effects or, with distances of 1 and 3 syllables, an effect in the predicted direction (a change of 28 ms, or 7%, for the immediate distance; 15 ms, or 4%, for the distance of 3 syllables). The results will be evaluated below in a statistical model.

Two LME models were created. First, using the addition of glottalization dataset, the manipulation condition and the distance of prompt from target were not significant ($\chi^2(1) = 2.1$, $p = 0.147$; $\chi^2(4) = 5.3$, $p = 0.253$). Likewise, there was no significant interaction between the two effects ($\chi^2(4) = 5.0$, $p = 0.287$). However, in the second model (the elimination dataset), there was a significant interaction of manipulation * distance ($\chi^2(4) = 21.9$, $p < 0.001$). Specifically, post-hoc tests revealed a significant effect in the immediate condition one syllable from the target (Linked – Glottalized = 0.1024, SE = 0.0237, t-ratio = 4.327, $p < 0.001$). After back transformation, such an effect corresponds to 40 ms (or a change of 11%). No other distances were associated with significantly different pairs ($p > 0.05$). The effect plots are shown in Figure 5.

4.3 Discussion

The aim of our study was to contribute to the research into the functionally-defined phenomenon of boundary glottalization, mostly in relation to its impact on the cognitive load of native Czech listeners. This is an extension to previous research (Bissiri et

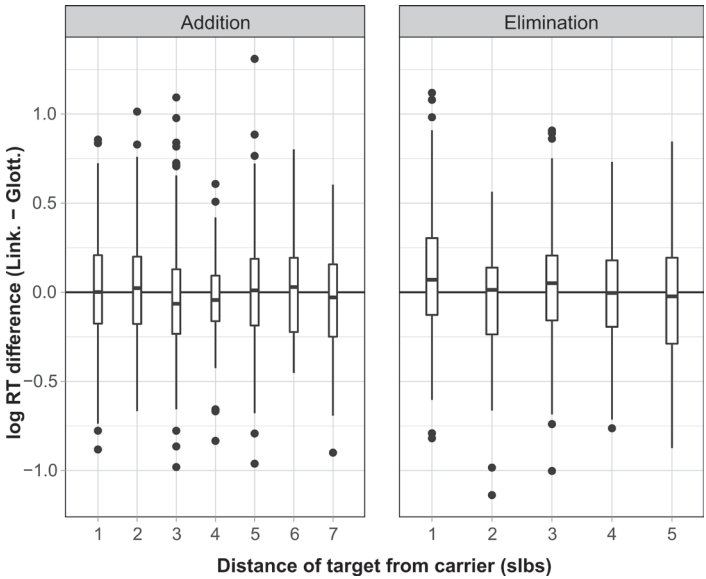

Figure 4. A difference in logarithmic RTs of glottalized and linked stimuli in two types of manipulation as a function of distance. Positive values indicate shorter RTs in the glottalized stimuli.

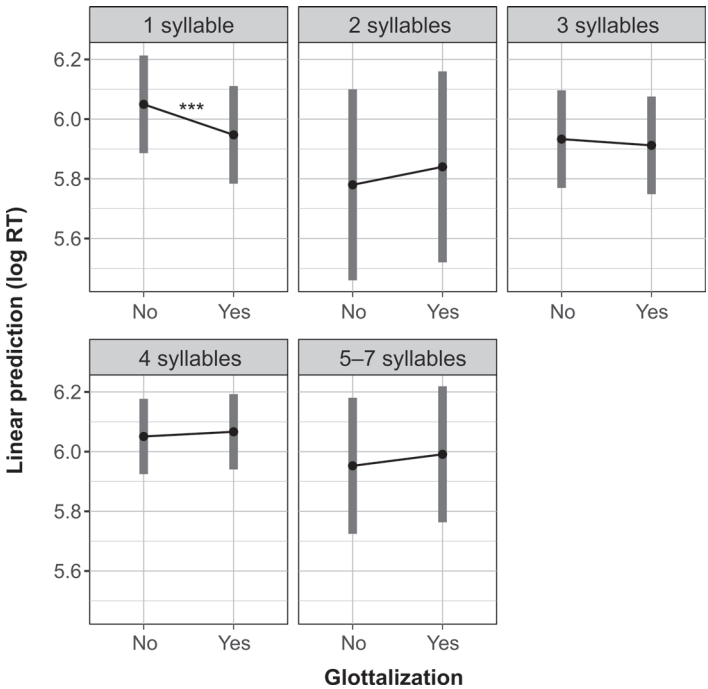

Figure 5. Effect plots of the second LME model (elimination of glottalization manipulation) with an interaction between glottalization and distance of prompt word from carrier word.

al., 2011; Volín et al., 2012; Schwartz et al., 2015), which used non-native listeners as participants. The hypothesis was that the presence of glottalization in the stimulus lowers the cognitive load, and thus speeds up the processing of the following words. To test this prediction, we measured reaction times (RTs) in a word-monitoring experiment. In contrast to the previous studies, we used a more spontaneous material (TV debate) and we also focused on the distance of the glottalization from the target word, which makes our research innovative.

The results of the behavioural experiment partly confirmed the prediction. Stimuli with glottalization were associated with shorter RTs than stimuli with linking, but only when the target word (the prompt with an initial consonant) appeared one syllable after the manipulated carrier context. The cognitive benefit of glottalization thus seems to be restricted to immediate surroundings, and is not extended to more distant words. This would be in line with the results of the previous studies, which used an even shorter distance (zero syllables: the prompt was identical to the carrier word). Moreover, the effect was apparent only in one of two types of manipulation (elimination of glottalization rather than its addition).

Interestingly, upon a reviewer's question regarding gender differences, we examined post hoc whether male and female participants behaved similarly with respect to the glottalization effect. Although male participants were somewhat faster, which is in accordance with previous research (e.g., Jain, Bansal, Kumar & Singh, 2015), we did not find any reliable differences in the performance of the two groups in relation to the presence of glottalization. Therefore, glottalization seems to be perceived similarly by both genders, in contrast to its production (usage).

However, the merit of our research cannot be reduced to the results only; what seems to be equally valuable are methodological observations from the design of the experiment. Although the use of material of a lower level of control has its justification in the attempt to approximate common communicative situations, our experiment has shown that employing such a type of material in a perception test offers several obstacles which may, without sufficient attention, reduce the validity of the experiment. One of the pitfalls is the stimuli selection process, which is restricted by the available recordings and the factors included in them. More attention and control should be paid to the key factors, such as balancing the number of stimuli with respect to prompt distance, stimulus length, or the semantic status of the carrier words. Such a process will place higher demands on the extent of available material and the researcher's time.

Another suggestion for future research relates to the creation of the manipulated stimuli. Such a process may disturb the temporal characteristics of the original recordings. It is a well-known fact that an interference with the rhythm of speech increases by itself the cognitive load of processing. Despite our attempts to minimize the temporal changes introduced by manipulation, it is important to bear in mind that manipulated and original items automatically have a different starting point, regardless of the level of the independent variable (linked vs. glottalized in our case). Results could then be biased or misinterpreted. In fact, the hypothesis that glottalization leads to shorter RTs should be supplemented with the hypothesis that manipulation leads to longer RTs. In this light, our results (and those of Bissiri et al., 2011 and Volín et al., 2012) make more sense. The addition of glottalization combines two opposing effects, which produced null or very

weak outcomes, whereas the elimination of glottalization combines two parallel effects, which yielded outcomes of different strength in the expected direction. Solution to this methodological problem is an important task for researchers in the field of RT measurement. An obvious escape route – presenting natural, non-manipulated stimuli, as in Schwartz et al. (2015) – does not seem to be the way to go. Their results showed only weak effects, which might be due to the fact that the paired versions were not completely identical. Only a single point of difference (glottalization) is clearly needed between the matched versions.

5. Conclusions

According to our hypothesis, the presence of word-initial glottalization should decrease the cognitive difficulty of sentence processing, leading to a decrease in reaction times (RTs) for glottalized items. This prediction was confirmed partially, as the effect was restricted to immediate contexts (distance of one syllable between the prompt word and the carrier word) and to one type of manipulation (elimination as opposed to addition of glottalization). The experiment with native listeners thus confirmed some of the previous results from L2 processing, and highlighted certain methodological aspects of an RT experiment design. Finally, the experiment showed both merits and difficulties of using speech material characterized by a lower level of control than the commonly used read speech.

Acknowledgements

This study was supported from the European Regional Development Fund-Project 'Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World' (No. CZ.02.1.01/0.0/0.0/16_019/0000734) and was created within the programme 'Cooperatio', scientific field Linguistics.

REFERENCES

Ashby, M. & Przedlacka, J. (2011). The stops that aren't. In: *English Phonetics 14–15. Festschrift commemorating the retirement and the 70th birthday of Professor John C. Wells*. The English Phonetic Society of Japan.

Ashby, M. & Przedlacka, J. (2014). Measuring incompleteness: Acoustic correlates of glottal articulations. *Journal of the International Phonetic Association*, 44(3), 283–296.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.

Bissiri, M. P., Lecumberri, M. L., Cooke, M. & Volín, J. (2011). The role of word-initial glottal stops in recognizing English words. In: *Proceedings of Interspeech 2011*, Florence, pp. 165–168.

Boersma, P. & Weenink, D. (2021). *Praat: Doing phonetics by computer (version 6.1)* [Computer software]. Retrieved from <http://www.praat.org/>.

Cruttenden, A. (2014). *Gimson's Pronunciation of English* (8th ed.). Taylor and Francis.

Di Napoli, J. (2015). Glottalization at phrase boundaries in Tuscan and Roman Italian. In: J. Romero & M. Riera (Eds.), *The Phonetics–Phonology Interface: Representations and Methodologies* (pp. 125–147). John Benjamins.
Dilley, L., Shattuck-Hufnagel, S. & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, 24, 423–444.
Docherty, G. & Foulkes, P. (1999). Sociophonetic variation in 'glottals' in Newcastle English. In: *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, pp. 1037–1040.
Fabricius, A. (2002). Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English World-Wide*, 23(1), 115–136.
Forster, K. I. & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, 35(1), 116–124.
Gavaldà, N. (2016). Individual variation in allophonic processes of /t/ in Standard Southern British English. *The International Journal of Speech, Language and the Law*, 23(1), 43–69.
Gobl, C. & Ní Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40(1–2), 189–212.
Gordon, M. & Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, 29, 383–406.
Grászi, T. E. & Markó, A. (2018). Word-initial glottal marking in Hungarian as a function of articulation rate and word class. In: *Challenges in Analysis and Processing of Spontaneous Speech* (pp. 75–98). mta Nyelvtudományi Intézet.
Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, 31(3–4), 423–445.
Hála, B. (1967). *Výslovnost češtiny I*. Academia.
Jain, A., Bansal, R., Kumar, A. & Singh, K. D. (2015). A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied and Basic Medical Research*, 5(2), 124–127.
Jiang, N. (2012). *Conducting Reaction Time Research in Second Language Studies*. Routledge.
Keating P., Garellek, M. & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. In: *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, paper 821.
Kilborn, K. & Moss, H. (1996). Word monitoring. *Language and Cognitive Processes*, 11(6), 689–694.
Kohler, K. (1994). Glottal stops and glottalization in German. *Phonetica*, 51, 38–51.
Kopečková, M. (2020). *Analýza zvukové roviny mluvního projevu moderátorů hlavního TV zpravodajství* [Ph.D. thesis]. Univerzita Palackého v Olomouci, Filozofická fakulta.
Ladefoged, P. & Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell.
Lenth, R. V. (2022) *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.7.2) [Computer software]. Retrieved from <https://CRAN.R-project.org/package=emmeans>.
Local, J., Wells, W. H. G. & Sebba, M. (1985). Phonology for conversation: Phonetic aspects of turn delimitation in London Jamaican. *Journal of Pragmatics*, 9, 309–330.
Maddieson, I. (1984). *Patterns of Sounds*. Cambridge University Press.
Malisz, Z., Żygis, M. & Pompino-Marschall, B. (2013). Rhythmic structure effects on glottalisation: A study of different speech styles in Polish and German. *Laboratory Phonology*, 4(1), 119–158.
Ogden, R. (2001). Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association*, 31(1), 139–152.
Olaya, N. P. (1967). *A Phonological Grammar of a Dialect of Ilokano* [M.A. thesis]. University of British Columbia.
Palková, Z. (2016). Hlasivkový ráz ve výzkumu zvukové stavby češtiny. In: A. Cychnerska & I. Sawicka, *Sandhi w językach słowiańskich II* (pp. 143–158). Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
Pompino-Marschall, B. & Żygis, M. (2010). Glottal marking of vowel-initial words in German. *ZAS Papers in Linguistics*, 52, 1–17.
R Core Team (2020). *R: A language and environment for statistical computing (Version 4.0.3)* [Computer software]. R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org>.
Redi, L. & Shattuck-Hufnagel, S. (2001). Variation in the realization of glottalization in normal speakers. *Journal of Phonetics*, 29, 407–429.

Schwartz, G., Rojczyk, A. & Balas, A. (2015). Monitoring English sandhi-linking: A study of Polish listeners' L2 perception. *Research in Language*, 13(1), 61–76.
Skákal, L. (2013). *Užívání hlasivkového rázu u rodilých a nerodilých mluvčích francouzštiny* [B.A. thesis]. Univerzita Karlova, Filozofická fakulta.
Skákal, L. (2015). *Užívání glotalizace jako faktor umožňující identifikaci mluvčího* [M.A. thesis]. Univerzita Karlova, Filozofická fakulta.
Skarnitzl, R. (2004). Acoustic categories of nonmodal phonation in the context of the Czech conjunction "a". In: Z. Palková & J. Veroňková (Eds.), *AUC Philologica 1/2004, Phonetica Pragensia X* (pp. 57–68). Karolinum.
Skarnitzl, R., Čermák, P., Šturm, P., Obstová, Z. & Hricsina, J. (2021). Glottalization and linking in the L2 speech of Czech learners of Spanish, Italian and Portuguese. *Second Language Research*, 38(4), 941–963. DOI: 10.1177/02676583211015803.
Šturm, P. & Bičan, A. (2021). *Slabika a její hranice v češtině*. Karolinum.
Šturm, P. & Volín, J. (2012). Měření reakčních dob u experimentů s akustickými podněty. *Akustické listy*, 18(2–4), 25–30.
Veroňková J. (2016). Výskyt rázu ve čtených a semispontánních projevech v češtině. In: A. Cychnerska & I. Sawicka, *Sandhi w językach słowiańskich II* (pp. 159–172). Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
Volín, J. (2012). Jak se v Čechách „rázuje". *Naše řeč*, 95(1), 51–54.
Volín, J., Uhrinová, M. & Skarnitzl, R. (2012). The effect of word-initial glottalization on word monitoring in Slovak speakers of English. *Research in Language*, 12(2), 173–181.
Wagner, P., Trouvain, J. & Zimmerer, F. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics*, 48, 1–12.
Wickham H. (2009). *Ggplot2: Elegant Graphics for Data Analysis*. Springer.
Wolk, L., Abdelli-Beruh, N. B. & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, 26(3), e111–e116.
Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85(3), 315–337.

Appendix

Target items used in the perception test. All caps indicate the prompt (target word) displayed on the screen. Carrier words in bold are sites of glottalization where manipulations were performed.

Originally glottalized items:	Originally linked items:
1) *by si měl, nebo si **uvědomil** ZŘEJMĚ, a přišel taky*	21) *taková by měla být **omezena** POUZE na projevy učiněné v souvislosti*
2) *je, tak by měl kopírovat **agendu** MINISTERSTVA kultury*	22) *poslanci by neměli mít **imunitu**, DISKUTOVAL bych o tom*
3) *to, že ti **občané** MUSÍ mít pocit*	23) *má být **ošetřováno** RŮZNĚ jízdné a podobně*
4) *najevo. Pan Milota skončil **okamžitě** V PODSTATĚ ve funkci*	24) *priority jsou tak **odlišné**, že KAŽDÝ ví, že*
5) *si **uvědomil**, že jeho MÍSTO je jinde a setsakramentsky*	25) *základní pilíře, prevence, průhlednost **a** POSTIH. Všechny veřejné*
6) *nikoliv **už** za MÍSTOPŘEDSEDU poslaneckého výboru*	26) *zástupci z řad **občanů**, to ZNAMENÁ bude vytvořen registr*
7) *Jsem přesvědčen **o** tom že PROSTĚ*	27) *se může přihlásit **a** BUDE přímo u zadávání*

Originally glottalized items:	Originally linked items:
8) *vás těším **opět** PŘÍŠTÍ neděli v jedenáct hodin*	28) *Bude moct **účinně** PROTI korupci bojovat*
9) *tu vládu **asi** víc TLAČIT, protože již předminule*	29) *pochybnost, že prostě nějaký **úředník** ve SKLEPENÍ magistrátu dělá něco*
10) *chystali do **opozice** po KVĚTNOVÝCH volbách sněmovních*	30) *pomalu, ta **otázka** ZNĚLA, jestli bude*
11) *my nebudeme nikoho kádrovat, **abychom** ŘÍKALI, že ten je vhodný*	31) *v další části budeme mluvit **o** BOJI proti korupci*
12) *nevšiml jsem si, že by **akciová** společnost FUTURA měla nějak*	32) *prezidentský kandidát **udělá** KAMPAŇ, jakou uzná*
13) *ten problém, **o** KTERÉM se tolik nemluví*	33) *v pátek jednání s **odboráři** DOPRAVNÍHO podniku*
14) *to měl být, mně přijde **úplně** LICHÉ, protože*	34) *několika dnů vyloučeni z **občanské** DEMOKRATICKÉ strany*
15) *každý **občan** ČESKÉ republiky má být právo*	35) *se to takhle říct **od** STOLU, protože já*
16) *to, **abychom** se jako POLITICKÉ strany dohodli*	36) *určitě ano, jak **imunita**, tak POSLANECKÉ náhrady*
17) *bych měl jít příkladem, **a** NEMŮŽEME od lidí něco chtít*	37) *je to například na **úřadech**, že JEDNOU za rok*
18) *ti kteří vlastně o těch zakázkách **a** MILIARDÁCH a směřování*	38) *něco jako **omezení** HORNÍ hranice*
19) *dnes pozice **imunity** ZTRATILA svůj význam*	39) *Že zástupce komunistické strany se **odvolává** na RYCHLÉ šípy a na na skauty*
20) *Dalších **institucí** je NÍZKÁ i z těchto důvodů*	40) *bych se ani nebavil **o** těch NÁKLADECH na kampaň, protože*

Michaela Svoboda
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
E-mail: michaela.svoboda@post.cz

Pavel Šturm
Institute of Phonetics
Faculty of Arts, Charles University
Prague, Czech Republic
E-mail: pavel.sturm@ff.cuni.cz